Published online 23 June 2009

Nucleic Acids Research, 2009, Vol. 37, No. 15 ***5093–5104***
doi:10.1093/nar/gkp532

A *trans*-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte *Isoetes engelmannii*

Felix Grewe[1], Prisca Viehoever[2], Bernd Weisshaar[2] and Volker Knoop[1,*]

[1]Institut für Zelluläre und Molekulare Botanik (IZMB), Universität Bonn, Kirschallee 1, 53115 Bonn and [2]Institut für Genomforschung und Systembiologie (IGS), Universität Bielefeld, Universitätsstraße 25, 33594 Bielefeld, Germany

Received March 19, 2009; Revised June 5, 2009; Accepted June 6, 2009

ABSTRACT

Plant mitochondrial genomes show much more evolutionary plasticity than those of animals. We analysed the first mitochondrial DNA (mtDNA) of a lycophyte, the quillwort *Isoetes engelmannii*, which is separated from seed plants by more than 350 million years of evolution. The *Isoetes* mtDNA is particularly rich in recombination events, and chloroplast as well as nuclear DNA inserts document the incorporation of foreign sequences already in this most ancestral vascular plant lineage. On the other hand, particularly small group II introns and short intergenic regions reveal a tendency of evolution towards a compact mitochondrial genome. RNA editing reaches extreme levels exceeding 100 pyrimidine exchanges in individual mRNAs and, hitherto unobserved in such frequency, also in tRNAs with 18 C-to-U conversions in the tRNA for proline. In total, some 1500 sites of RNA editing can be expected for the *Isoetes* mitochondrial transcriptome. As a unique molecular novelty, the *Isoetes cox*1 gene requires *trans*-splicing via a discontinuous group I intron demonstrating disrupted, but functional, RNAs for yet another class of natural ribozymes.

INTRODUCTION

Mitochondrial DNAs (mtDNAs) trace back in evolution to the genome of an α-proteobacterial endosymbiont which gave rise to the mitochondria of eukaryotic cells (1). The mitochondrial genomes in most animal (metazoa) lineages are compact, circular DNAs of some 16 kb which encode a standard set of 37 or fewer tightly packed genes (2). The mtDNAs of other eukaryotes, however, are significantly more diversified, most notably between different protist lineages, which reflect most of the evolutionary history and diversity of eukaryotic cells (3). These, for example, include obvious evolutionary ancestral states such as the gene-rich 69-kb mtDNA of the jakobid protist *Reclinomonas americana* with nearly 100 mitochondrial genes (4) as well as the massively reduced 6-kb mtDNA of the malaria parasite *Plasmodium falciparum* with only five genes (5), reflecting a massive gene transfer into the nuclear genome.

Land plant (embryophyte) mtDNAs in contrast are significantly extended in size and may exceed 2000 kb in certain flowering plant (angiosperm) families (6). The embryophyte mtDNAs encode some of the genes for protein subunits of the respiratory chain complexes, for ribosomal proteins and for proteins involved in cytochrome *c* biogenesis which are found in protists but are generally absent from animal or fungal mitochondrial genomes. Many plant mitochondrial genes are interrupted by introns belonging to either of the two classes of ribozyme-type group I or group II introns, which are commonly encountered in fungal, algal and plant organelle genomes and occasionally also in bacteria, phages and exceptionally also in the mtDNAs of primitive metazoan lineages (2,7). Besides intron gains, size increases of plant mitochondrial genomes have occurred mainly through the extension of non-coding intergenic regions. This becomes immediately apparent when the available mtDNAs of land plants (Figure 1) are compared to those of the charophyte algae (8,9) phylogenetically related to the embryophyte lineage (10). Some of the additional sequences, at least in flowering plant (angiosperm) mtDNA, have been identified as copies of chloroplast or nuclear DNA (11,12) or even as gained via horizontal gene transfer (13–16).

Moreover, plant mitochondria have evolved complex features with respect to genome arrangements and gene expression that are contrary to the general evolutionary trend of compaction and streamlining of endosymbiotic genomes (17,18). The gain of RNA editing activity in the organelles (19,20) to correct gene sequences by

*To whom correspondence should be addressed. Tel: +49 228 73 6466; Fax: +49 228 73 6467; Email: volker.knoop@uni-bonn.de

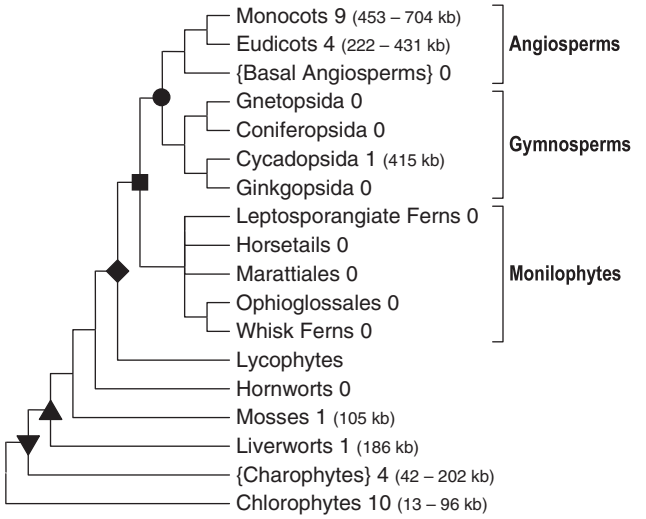

Figure 1. Current view of a simplified phylogeny of extant Viridiplantae (green plants *sensu lato*). The cladogram shown summarizes insights from recent molecular studies of land plant phylogeny (e.g. Qiu *et al.*, 2006). Numbers of completely sequenced mtDNAs (http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/plants_tax.html) are indicated for each group. Brackets indicate paraphyletic grades, all other designations indicate reasonably well supported monophyletic groups. Further well supported monophyletic clades of higher order are the spermatophytes (seed plants, circle), the euphyllophytes (square), the tracheophytes (vascular plants, rhomb), the embryophytes (land plants, up triangle) and the streptophytes (down triangle) whereas the bryophytes (liverworts, mosses and hornworts) are paraphyletic.

pyrimidine exchanges (mainly cytidine to uridine) at the transcript level likewise appears to be gained with the earliest embryophytes, although this phenomenon is suspiciously absent in the subclass of marchantiid liverworts (21). Despite size increase to more than 100 kb in early embryophyte evolution (Figure 1), the two so far available completely sequenced mtDNAs of bryophytes—those of the liverwort *Marchantia polymorpha* (22) and the moss *Physcomitrella patens* (23)—are recognized as simple, circular-mapping genomes. However, linear DNAs may in fact contribute significantly to the population of mtDNA molecules actually present in the mitochondria (24). Flowering plant (angiosperm) mtDNAs are rich in active recombination resulting in co-existing alternative mitochondrial genome arrangements (e.g. 25–27), the stoichiometries of which are now understood to be regulated by nuclear-encoded protein factors related to bacterial *rec* proteins (28). A 'master-circle' representing the full mitochondrial genome complexity in a single circular DNA molecule may be entirely hypothetical in these cases (29).

Evidently coinciding with the rise of recombinational activity during the evolution of plant mitochondrial genomes is the appearance of *trans*-splicing group II introns producing peculiar arrangements of the affected genes with exons distributed across wide distances in the mtDNA. The origins of *trans*-splicing group II introns have been traced back through plant evolution (Figure 1) as having arisen through disruption of ancestral, conventional group II introns that can still be identified as their respective orthologues in ferns, hornworts and mosses (30,31).

In the absence of complete mtDNA information for ferns, horsetails, lycophytes or hornworts (Figure 1), there is currently a large phylogenetic gap remaining between the available mtDNA sequences of the liverwort *Marchantia* or the moss *Physcomitrella* and the first recently completed mtDNA of a gymnosperm, the cycad *Cycas taitungensis* (32). Accordingly, we have investigated the mtDNA of the quillwort *Isoetes engelmannii*. As a lycophyte, *Isoetes* represents the most ancestral lineage of recent vascular plants (tracheophytes). The mtDNA of *I. engelmannii* offers a plethora of surprising findings, which include particularly small group II introns, extreme frequencies of DNA recombination and RNA editing also in tRNAs, insertions of chloroplast and nuclear DNA and, most notably, a *trans*-splicing group I intron.

MATERIALS AND METHODS

Fosmid analyses

Isoetes engelmannii plant material originally collected in South Central Indiana (USA) by Jerry Gastony, and subsequently greenhouse cultivated, was kindly made available through Jeff Palmer and Erin Badenhop (Bloomington, IN). The non-green bulb tissue of plants was used to enrich for mitochondrial vs. chloroplast DNA. Total genomic DNA was isolated using a CTAB protocol. After size-fractionation into ~38 kb fragments, DNA was blunt-ended and cloned into the fosmid vector pCC1FOS using the CopyControl Fosmid Library Production Kit (EPICENTRE, Madison, Wisconsin). A library of 11 700 fosmid clones was sorted and filter-spotted for successive rounds of hybridization initially using a mixture of PCR-derived gene probes of *cox*3, *nad*2, *nad*5 and *nad*7 and subsequently with probes derived from the sequence-verified mitochondrial fosmids. Identity of fosmid clones was initially verified through terminal insert sequencing and positive clones were used for sub-library production. Fosmid DNAs were isolated using NucleoBond Xtra Midi EF Kit (Macherey Nagel, Düren, Germany), sheared by Nebulizers (Invitrogen, Carlsbad, California), blunted using a End-It DNA End-Repair Kit (EPICENTRE, Madison, Wisconsin), A-tailed with Taq-Polymerase (Genaxxon, Biberach, Germany), and fractionated by preparative electrophoresis in 0.8% agarose. Fragments of 2–2.5 kb in size were recovered using the NucleoSpin Extract II Kit (Macherey Nagel, Düren, Germany) and cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin). Minimally 400 plasmid clones were sequenced for each fosmid to reach ~8-fold sequencing coverage. Five fosmid clones (11P20, 19N12, 26A6, 28M14 and 30K18) were validated as native mtDNA. Graphical maps of the fosmid clones created with OGDRAW v1.1 (33) are given in Supplementary Figure 1. The respective fosmid insert sequences were annotated and deposited in the database under accession numbers FJ010859, FJ536259, FJ390841, FJ176330 and FJ628360, respectively.

Sequence analyses

Sequence handling and analysis of final fosmid assemblies was essentially done using the alignment explorer of the MEGA software (34). Identification of loci was essentially done using similarity searches with Basic Local Alignment Search Tool (BLAST) service at the NCBI (35). Candidate sites of RNA editing in *Isoetes* were identified manually in alignments of deduced protein sequences with homologues in *Chara* (AY267353) and *Marchantia* (M68 929), species devoid of RNA editing.

Trancript analyses

Total *I. engelmannii* RNA was isolated using the NucleoSpin RNA Plant Kit (Macherey Nagel, Düren, Germany); cDNA was synthesized with RevertAid First Strand cDNA Synthesis Kit (Fermentas, Burlington, Ontario) in the presence of random hexamer primers as recommended by the manufacturer. Oligonucleotide pairs (all sequence information available from the authors upon request) were used for RT–PCR amplification according to the standard protocol of GoTaq DNA Polymerase (Promega, Madison, Wisconsin) in a GeneAmp PCR System 2700 (Applied Biosystems, Foster City, California) with annealing temperatures between 50°C and 55°C. Amplicons were recovered from agarose gel and cloned into pGEM T Easy vector as described above. On average, 10 cDNA clones per locus were sequenced and analysed by comparison with the corresponding DNA sequences. RNA self ligation for *cox*1 transcript end mapping followed published procedures (36). Total *I. engelmannii* RNA was ligated by T4 RNA ligase (New England Biolabs, Ipswich, Massachusetts) and cDNA was synthesized with RevertAid First Strand cDNA Synthesis Kit (Fermentas, Burlington, Ontario) in the presence of 200 pmol of outward directed primers cox1leftdo1 and cox1rightup (1 and 2, respectively, in Figure 2). The same oligonucleotide pairs were used for first PCR amplification according to the standard protocol of BD Advantage 2 polymerase (BD Bioscience, Franklin Lakes, New Jersey) in a GeneAmp PCR System 2700 (Applied Biosystems, Foster City, California) with annealing temperatures at 45°C. Amplification results in several molecules of different sizes which were recovered from an agarose gel using the NucleoSpin Extract II Kit (Macherey Nagel, Düren, Germany). These molecules served as templates for subsequent nested PCR in presence of primers cox1leftdo and cox1rightup (1 and 3, respectively, in Figure 2) with same preferences as initial PCR.

Southern blotting

For Southern blotting, ~10 µg of *I. engelmannii* total DNA was digested with combinations of restriction enzymes (either *Eco*RI and *Cfr*9I or *Eco*RI and *Eco*RV, see Figure 6) and separated on a 0.8% agarose gel prior to blotting following established procedures (37). Approximately 100 ng of PCR-derived probes (Figure 6) were radioactively labelled with 50 µCi of α-P^{32}-dCTP. Hybridization of the nylon blotting membranes was overnight at 65°C in 50 mM sodium phosphate buffer containing 0.9 M NaCl, followed by washing in 2× SSC with 0.1% SDS at 65°C before exposure on a phosphor imager.

RESULTS

Genomic features and gene complement of the *Isoetes* mtDNA

The *I. engelmannii* mtDNA sequence was assembled from fosmid clones, identified in an arrayed library by hybridization with mitochondrial gene probes and verified in their mitochondrial nature through complete sequencing of the inserts. As more fosmid sequences were analyzed in the course of our studies, it became apparent that the same genes were repeatedly identified. Mitochondrial genes were found in different genomic environments, indicating a particularly high frequency of recombination events resulting in co-existing alternative gene arrangements (Figure 2). A total of 24 recombination breakpoints were identified, making the physical existence of a potential mtDNA master-circle encompassing the full mtDNA complexity highly unlikely. Different fosmid inserts reflected different pathways through the recombination points and the resulting products of DNA recombination were exemplarily verified as co-existing (see below). The net mtDNA sequence complexity of the analyzed *I. engelmannii* fosmid clones is 57 571 bp, with an overall A + T content of 51.3% and a percentage of 46.2% coding sequences.

We identified a typical complement of plant mitochondrial genes (Table 1) encoding subunits of respiratory chain complex I (*nad* genes *nad*1, 2, 3, 4, 4L, 5, 6, 7 and 9), complex II (*sdh*3), complex III (*cob*), complex IV (*cox*1, 2 and 3) and of complex V, the ATP synthase (*atp*1, 4, 6, 8 and 9). Likewise present are the genes for the large, small and 5S rRNAs (*rrn*L, *rrn*S, *rrn*5), for four ribosomal proteins (*rpl*5, *rps*2, *rps*3 and *rps*4) as well as the *tat*C gene encoding a subunit of the *sec*-independent transport pathway, and thirteen intact tRNA genes. Hence, on the one hand, four ribosomal protein genes demonstrated to be frequently transferred to the nucleus in angiosperms (38) are present in the *Isoetes* mtDNA. On the other hand, genes encoding cytochrome biogenesis components (*ccm*B, *ccm*C, *ccm*F) are completely lacking as had previously been observed for the land plant lineage only in the mtDNA of the green alga *Chaetosphaeridium* (8). To exclude the possibility that the *ccm* genes were accidently missed through yet a further recombination event, we have used oligonucleotide primers directed against conserved *ccm*B, *ccm*C and *ccm*F sequences but were unable to retrieve them in PCR approaches using *I. engelmannii* DNA. In addition, we identified small pseudogene fragments of three tRNA genes, of the *rrn* genes and of the *nad*4 and *rps*1 genes.

Recombination points (Figure 2) were identified both in intergenic regions (R1, R2, R8, R11, R20 and R21) as well as in coding regions (of *atp*6: R5, *atp*8: R12, *atp*9: R19, *cob*: R3, R22, *cox*1: R12, R13, R4, R14, *rpl*5: R10, *rps*2: R23, *rps*3: R4, *trn*F: R18, *nad*1: R6, R10, *nad*2: R24, R9, R15, *nad*3: R17, *nad*5: R16, R3b and *sdh*3: R7,

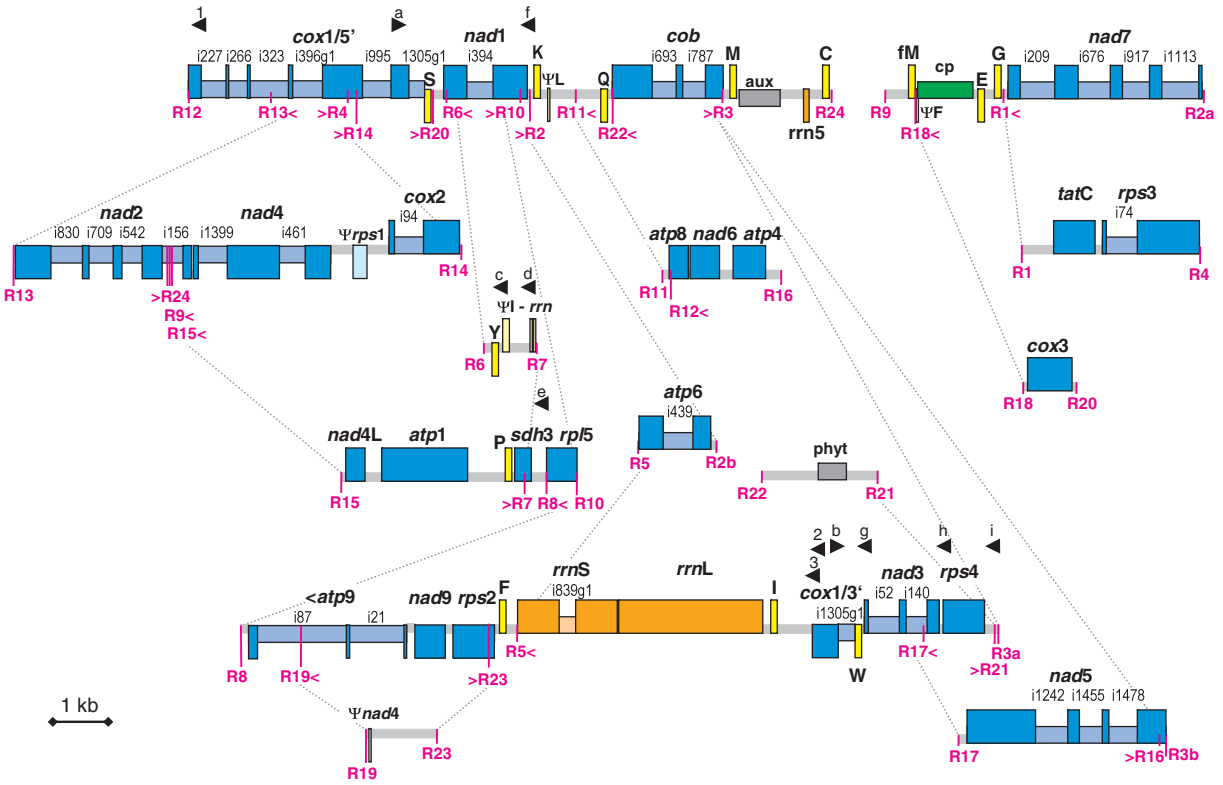

Figure 2. The *Isoetes engelmannii* mtDNA with protein-coding genes shown in blue, tRNAs in yellow, rRNAs in orange, pseudogenes in the respective lighter colours and the cp and nuc DNA inserts indicated with green and grey boxes, respectively. Drawing is approximately to scale. Genes shown above or below the lines indicate directions of transcription to the right or to the left, respectively. Recombination points (R1–R24) are highlighted in magenta with arrows indicating recombination forks. Selected connections between islands of recombination are exemplarily shown with stippled grey lines. Unique net mtDNA sequences add up to 57 571 bp. Arrowheads indicate oligonucleotide primers (1–3 and a–i) anchoring in regions not affected by RNA editing to analyze the arrangement of *cox*1 and its transcript maturation.

respectively), which accordingly result in fragmented pseudogene copies co-existing with the functional genes. In agreement with the observation of highly frequent recombination, only two evolutionary ancient gene linkages (*trn*P–*sdh*3 and *nad*4–*nad*2) are conserved as mere traces of much more extended syntenies that were identified when the liverwort *Marchantia* and the moss *Physcomitrella* mtDNAs were compared (23) to those of the streptophyte algae *Chaetosphaeridium globosum* (8) and *Chara vulgaris* (9). When not affected by recombination, intergenic regions between functional genes are small (only 17 bp between *nad*4 and *nad*2 and 8 bp between *atp*8 and *nad*6, respectively) with the exception of the spacer between *trn*K and *trn*Q carrying the pseudo-*trn*L fragment and the non-coding regions extended through the insertions of foreign DNA fragments.

Insertions of foreign DNA

Three 'promiscuous' DNA inserts of foreign origin were identified in the *I. engelmannii* mtDNA. A 1208 bp fragment of chloroplast DNA located between *trn*E and *trn*fM (Figure 2) covers parts of the chloroplast *trn*A and 23S rRNA genes. Highest similarity of this chloroplast sequence insert is found with the corresponding chloroplast sequence of another *Isoetes* species deposited in the database (*I. malinverniana*, DQ629281) indicating (recent) inter-organellar rather than horizontal gene transfer. Sequence deviations of the chloroplast insert in the *I. engelmannii* mtDNA from the native chloroplast homologue show striking pattern of degeneration with only two base changes within 1080 nt of 23S rRNA but indels of exclusively 5 or 6 bp (Supplementary Figure 2).

A 735 bp sequence stretch in the intergenic region between *rrn*5 and *trn*M bears strong similarity with nuclear encoded auxin-responsive transcription factors and a 533-bp sequence with similarity to phytochrome genes occurs between *rps*4 and *trn*Q (Figure 2). Like the chloroplast insertion, both nuclear sequence inserts are non-functional pseudogene fragments.

RNA editing in mRNAs

The protein-encoding genes in the *I. engelmannii* mtDNA show a very strong requirement for mRNA editing via pyrimidine exchanges to reconstitute evolutionary conserved codons. Altogether more than 1420 positions (over 1200 C-to-U and 220 U-to-C changes) in the *Isoetes* mitochondrial transcriptome appear to be subject to editing (Supplementary Table 1). This includes the reconstitution of appropriate AUG start codons from ACG threonine codons which is required in 12 cases and

Table 1. The gene and intron complement of *Isoetes engelmannii* (Ie) mtDNA in comparison to the mtDNAs of the liverwort *Marchantia polymorpha* (Mp) the moss *Physcomitrella patens* (Pp), the gymnosperm *Cycas taitungensis* (Ct) and the angiosperm *Arabidopsis thaliana* (At)

Genes/introns	*Mp*	*Pp*	*Ie*	*Ct*	*At*
atp1	+	+	+	+	+
atp1i989g2	+ +				
atp1i1050g2	+ +				
atp1i1128g2		+			
atp4	+	+	+	+	+
atp6	+	+	+	+	+
atp6i80g2		+			
atp6i439g2			+		
atp8	+	+	+	+	+
atp9	+	+	+	+	+
atp9i21g2		+	+		
atp9i87g2	+ +	+	+		
atp9i95g2		+			
ccmB	+	+		+	+
ccmC	+	+		+	+
ccmF	+	+		+	+
ccmFCi829g2		+		+	+
cob	+	+	+	+	+
cobi372g2	+				
cobi420g1		+			
cobi693g2			+		
cobi783g2	+				
cobi787g2			+		
cobi824g2	+ +				
cox1	+	+	+	+	+
cox1i44g2	+ +				
cox1i178g2	+ +				
cox1i227g2			+		
cox1i266g2			+		
cox1i323g2			+		
cox1i375g1	+				
cox1i395g1	+ +		+		
cox1i511g2	+ +	+			
cox1i624g1	+	+			
cox1i730g1	+ +				
cox1i732g2		+ +			
cox1i995g2			+		
cox1i1064g2		+			
cox1i1116g1	+ +				
cox1i1305g1	+		*trans*		
cox2	+	+	+	+	+
cox2i94g2			+		
cox2i97g2	+				
cox2i104g2		+			
cox2i250g2	+ +				
cox2i373g2		+		+	
cox2i691g2		+		+	+
cox3	+	+	+	+	+
cox3i171g2	+				
cox3i506g2		+			
cox3i625g2	+				
nad1	+	+	+	+	+
nad1i287g2		+			
nad1i394g2			+	*trans*	*trans*
nad1i477g2				+	+
nad1i669g2				*trans*	*trans*
nad1i728g2		+		+ +	+ +
nad2	+	+	+	+	+
nad2i156g2		+	+	+	+
nad2i542g2			+	*trans*	*trans*
nad2i709g2	+		+	+	+
nad2i830g2			+		
nad2i1282g2				+	+
nad3	+	+	+	+	+
nad3i52g2			+		
nad3i140g2	+		+		

(continued)

Table 1. Continued

Genes/introns	*Mp*	*Pp*	*Ie*	*Ct*	*At*
nad4	+	+	+	+	+
nad4i461g2		+	+	+	+
nad4i548g2	+				
nad4i976g2				+	+
nad4i1399g2			+	+	+
nad4L	+	+	+	+	+
nad4Li100g2	+				
nad4Li283g2	+	+			
nad5	+	+	+	+	+
nad5i230g2		+		+	+
nad5i753g1	+	+			
nad5i1242g2			+		
nad5i1455g2		+ +	+	*trans*	*trans*
nad5i1477g2			+	*trans*	*trans*
nad5i1872g2				+	+
nad6	+	+	+	+	+
nad7	Ψ	+	+	+	+
nad7i140g2		+			+
nad7i209g2		+	+	+	+
nad7i336g2	+				
nad7i676g2			+	+	+
nad7i917g2			+	+	+
nad7i1113g2	+		+	+	
nad9	+	+	+	+	+
nad9i283g2		+			
rpl2	+	+		+	+
rpl2i28g2	+				
rpl2i917g2				+	+
rpl5	+	+	+	+	+
rpl6	+	+			
rpl16	+	+		+	+
rps1	+	+	Ψ	+	
rps2	+	+	+	+	
rps3	+	+	+	+	+
rps3i74g2			+	+	+
rps3i257g2				+ +	
rps4	+	+	+	+	+
rps7	+	+		+	+
rps8	+				
rps10	+			+	
rps10i235g2				+	
rps11	+	+		+	
rps12	+	+		+	+
rps13	+	+		+	
rps14	+	+		+	Ψ
rps14i114g2	+				
rps19	+	+		+	Ψ
rrn5	+	+	+	+	+
rrnL	+	+	+	+	+
rrnLi827g2	+				
rrnS	+	+	+	+	+
rrnSi839g1			+		
rrnSi1065g2	+ +				
sdh3	+	+	+	+	
sdh3i100g2		+			
sdh4	+	+			Ψ
tatC	+	+	+	+	+
trnA(ugc)	+	+			
trnC(gca)	+	+	+	+	+
trnD(guc)	+	+		+	
trnD(guc) cp					+
trnE(uuc)	+	+	+	+	+
trnF(gaa)	+	+	+	+	
trnG(gcc)	+	+	+	+	+
trnG(ucc)	+	+			
trnH(gug)	+	+			
trnH(gug) cp				+	+
trnI(cau)	+	+	+	+	+
trnK(uuu)	+	+	+	+	+
trnL(caa)	+	+		+	

(continued)

Table 1. Continued

Genes/introns	*Mp*	*Pp*	*Ie*	*Ct*	*At*
trnL(uag)	+	+	Ψ	+	
trnL(uaa)	+	+			
trnM(cau)	+	+	+		
trnM(cau) cp				+	+
trnMf(cau)	+	+	+	+	+
trnN(guu)	+			+	
trnN(guu) cp					+
trnP(ugg)	+	+	+	+	+
trnP(ugg) cp					Ψ
trnQ(uug)	+	+	+	+	+
trnR(acg)	+	+			
trnR(ucg)	+				
trnR(ucu)	+	+		+	
trnS(gcu)	+			+	+
trnSi43g2	+				
trnS(uga)	+	+	+	+	+
trnS(gga) cp				+	+
trnT(ggu)	+	+			
trnV(uac)	+	+			
trnV(uac) cp				+	
trnVi39g2 cp				+	
trnW(cca)	+	+	+	+	
trnW(cca) cp					+
trnY(gua)	+	+	+	+	+

Plus signs indicate presence of a gene or intron, respectively and Ψ indicates recognizable pseudogene remnants (indicated only where not accompanied by a co-existing functional gene copy). Intron designations are according to the position of the preceding nucleotide in the uninterrupted coding sequence in the reading frames of *M. polymorpha* with the addendum g1 or g2 indicating group I or group II introns, respectively. Double plus signs (++) indicate presence of intron-borne ORFs.

the introduction of stop codons which is required for nine reading frames, respectively. In fact, the introduction of both the start and stop codons at the same time to correctly define the reading frames is necessary in five genes: *atp*6, *atp*9, *cox*3, *nad*4 and *nad*4L. Reverse U-to-C editing is required in the majority of mRNAs to convert genomically encoded stop codons into conserved glutamine or arginine codons. Indeed, only three out of the 24 protein encoding genes in the *Isoetes* mtDNA (*atp*4, *nad*3, *nad*4L) are without any stop codons on DNA level. To confirm the expectations on RNA editing, we performed exemplary cDNA analyses (Supplementary Table 1). RNA editing was indeed found to affect more than one out of five amino acid identities in the *atp*1 reading frame (115 of 515 codons) through 105 C-to-U exchanges and 23 U-to-C exchanges in the mRNA, including a change of six codons in a row with the sequence motif H-C-STOP-T-P-S changed into Y-R-Q-M-S-L by four C-to-U and two U-to-C exchanges in the transcript (not shown). A full 14 stop codons are removed through U-to-C RNA editing in the case of the *atp*1 mRNA. Approximately one out of seven nucleotides is affected by RNA editing in the *atp*9 reading frame (Figure 3A) resulting in sense changes in one out of three codons exactly as predicted to reconstitute evolutionary conserved codons, including introduction of both the start and the stop codon. Typically, the steady state pool of flowering plant mitochondrial

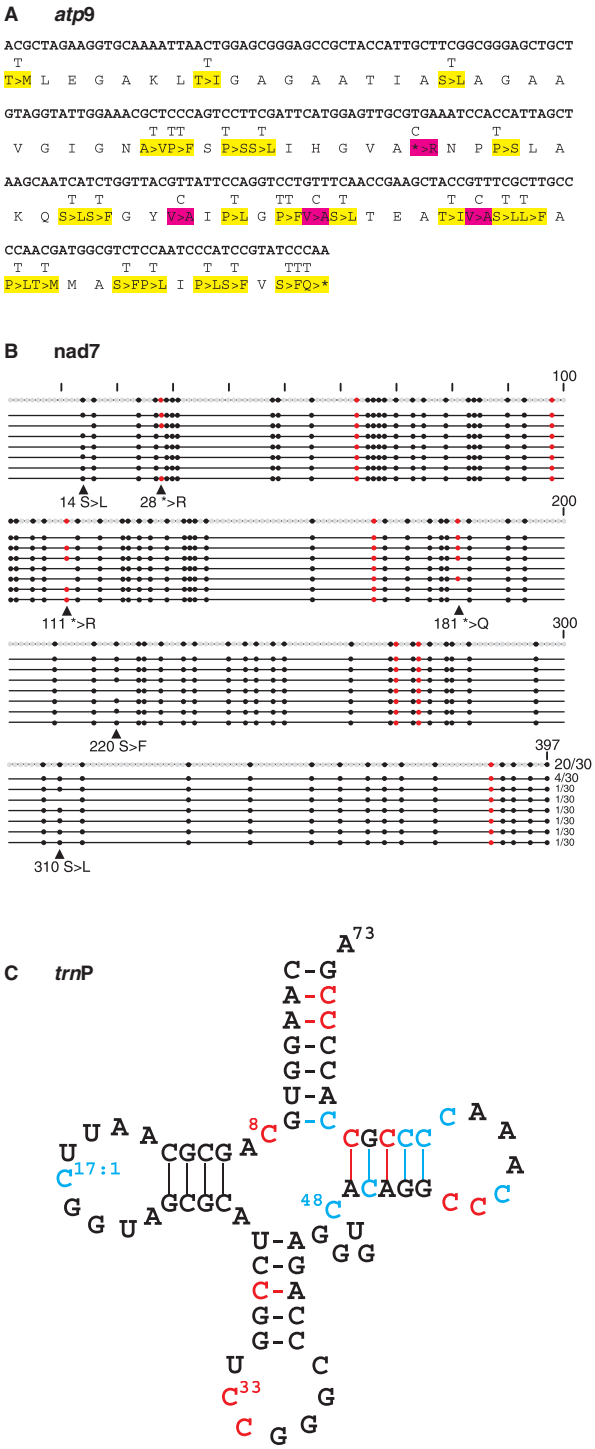

Figure 3. Examples of cDNA analyses for RNA editing: *atp*9 (A), *nad*7 (B) and *trn*P (C). (**A**) Yellow and magenta shading indicate changes introduced through C-to-U or U-to-C editing events in the *atp*9 mRNA. (**B**) The degree of partial editing of *nad*7 transcripts was investigated with 30 cDNA clones covering the complete reading frame of 397 codons. Black dots indicate codon sense changes, red dots the removal of stop codons. Twenty of 30 cDNA clones reflected complete editing at all predicted sites, the remaining clones lacked editing at certain positions. (**C**) Ten expected (red) and eight additional, non-predicted RNA editing sites (cyan) were found in *trn*P. Numbering of tRNA positions follows the standard convention with numbers after the colons indicating optional nucleotides not present in all tRNAs.

mRNAs contains transcripts edited to different degrees, reflecting only partial editing of some sites. To investigate this for *I. engelmannii* we examined 30 cDNA clones for the *nad*7 gene, for which we postulated RNA editing to correct 92 codon identities, including removal of nine stop codons. Complete editing of all the sites exactly as predicted was observed in 20 of the 30 cDNA clones, whereas four cDNAs lacked one editing to remove one of the stop codons (Figure 3B). The remaining six cDNA clones showed individual patterns lacking editing at this or another of one of five codons in total affected by partial editing (Figure 3B).

RNA editing in tRNAs

Cloverleaf modelling of the 13 tRNAs present in the *Isoetes* mtDNA strongly suggested frequent RNA editing activity to act on tRNAs as well. Several base-pairings in the four conserved stems and unpaired conserved uridines need to be re-established through C-to-U RNA editing in eleven tRNAs. The number of sites with predicted RNA editing events varied from single positions each in tRNA-fM and tRNA-G to six in tRNA-Q and even ten in tRNA-P (Figure 3C), respectively. Assuming that tRNA editing may take place in a precursor-transcript before processing we targeted a likely co-transcript of *trn*P with *sdh*3 (Figure 2) by RT–PCR, one of the rare cases of an ancient, conserved gene arrangement. Sequencing the cDNA product revealed not only the ten postulated positions of C-to-U exchanges, but also eight additional sites of C-to-U editing (but no reverse U-to-C changes), i.e. a total of 18 RNA editing positions in the *trn*P coding sequence (Figure 3C).

Introns

A total of 27 group II introns were identified in the *I. engelmannii* mtDNA (Table 1, nomenclature according to ref. (39), all of which are located in protein coding genes and most of which are particularly small. In fact, the *I. engelmannii* intron cox1i266 has a size of only 327 bp—to our knowledge, the smallest known group II intron as yet identified in any organism. Despite the strong recombinational activity in *I. engelmannii* mtDNA, none of the group II introns is in a *trans*-splicing arrangement. On the contrary, four of the known *trans*-splicing group II introns in angiosperms have *cis*-arranged counterparts in *I. engelmannii* (nad1i394, nad2i542, nad5i1455 and nad5i1477), seed plant introns nad1i669 and nad1i728 (in e.g. *Beta vulgaris* and *Oryza sativa*) obviously only appear later in evolution and get disrupted into *trans*-arrangements. A total of nine group II introns appear at novel insertion sites not yet observed in green algae (Charophytes or Chlorophytes), bryophytes or seed plants: atp6i439, cobi693, cobi787, cox1i227, cox1i266, cox1i323, cox1i995, cox2i94 and nad2i830 (Table 1).

Three group I introns were found in the *Isoetes* mtDNA, one in *rrn*S and two in *cox*1 (Figure 2). Orthologues of group I intron cox1i395 had previously been identified the liverwort *Marchantia* and in the alga *Chaetosphaeridium*, both of which carry endonuclease ORFs typical for this intron class. The *I. engelmannii* counterpart now identified is a small group I intron of only 328 bp without an ORF and hence similarly size-reduced as the group II introns. Yet smaller with a size of only 237 bp is a group I intron (rrnSi839g1) in the small ribosomal RNA gene *rrn*S.

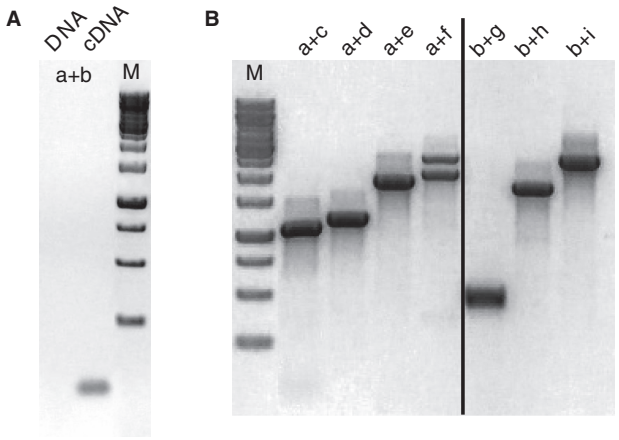

Figure 4. (**A**) Primers anchoring in the sub-terminal (a) and terminal (b) exons of *cox*1 fail to detect a genomic continuity across cox1i1305 or an alternative, intron-less *cox*1 copy with DNA but readily amplify the expected, spliced product with cDNA. (**B**) Each of the primers a and b combined with other primers anchoring in distant mtDNA regions (see Figure 2) downstream of *trn*Y (a + c), downstream of Ψ*trn*I (a + d), downstream of *sdh*3 (a + e), downstream of *nad*1 (a + f) or upstream of *trn*W (b + g), downstream of *nad*3 (b + h) or downstream of *rps*4 (b + i), respectively, all reveal amplicon products as expected. Two PCR products obtained simultaneously with primer combination a–f (lane 4) faithfully reflect the co-existing alternative genomic arrangements downstream of cox1/5′ to *nad*1 either via the *trn*S-*nad*1 continuity directly or alternatively through the *trn*S-R6-*trn*Y-Ψ*trn*I-R7-*sdh*3-*rpl*5-R10-*nad*1 pathway.

The most notable genomic peculiarity of the *Isoetes* mtDNA resides in the 3′ part of the *cox*1 gene (Figure 2). A group I intron (coxi1305) with a known homologue in the *Marchantia* mtDNA interrupts the *cox*1 coding region. However, intron homology breaks off sharply 210 bp after the splice donor site, 130 bp upstream of the *trn*S(uga) gene located downstream in inverted orientation. The seemingly missing terminal *cox*1 coding sequence was found elsewhere, preceded by the adequate splice acceptor site for joining the exons appropriately, 290 bp downstream of the *trn*W(cca) gene. Both parts of the *cox*1 coding regions have comparable similarities to other *cox*1 sequences in the database making independent, foreign origins, of the one or the other part of the gene, for example through horizontal gene transfer (15,16), unlikely.

To elucidate whether we had failed to identify a *cox*1 sequence continuity, we used primers anchoring in the directly flanking and also in distant *cox*1 exons, respectively, for PCR amplification assays on *I. engelmannii* DNA but failed to retrieve products (Figure 4A). To exclude potential malfunctions of the primers we used them individually in combinations with other primers anchoring in genomic distances ~2 kb apart in each case, as predicted from the recombinational mtDNA map (Figure 2). Expected products were retrieved

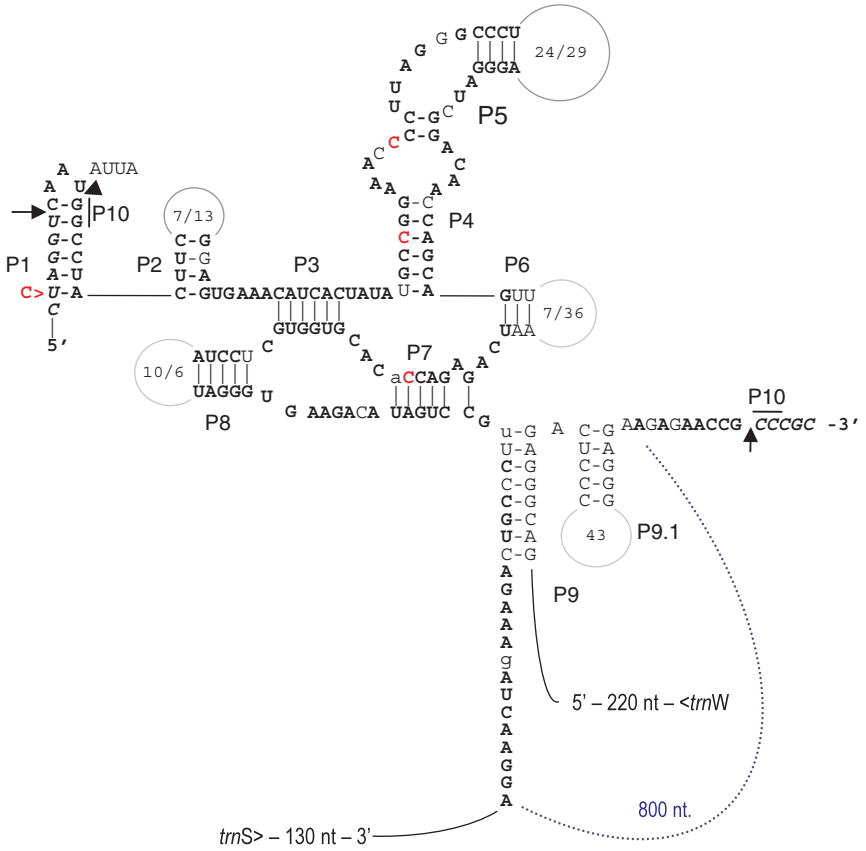

Figure 5. The *trans*-splicing group I intron cox1i1305 in *I. engelmannii*. The ribozymic core in the 5′-half (paired regions P1–P8) is conserved with the *cis*-arranged orthologue in *Marchantia polymorpha*. Italic letters indicate flanking exon sequences, bold letters indicate nucleotide identities in *Marchantia*, non-bold letters indicate transitions, lower case letters indicate transversions and numbers indicate loop sizes L2, L5, L6 and L8 in *Isoetes* and *Marchantia* (after the oblique), respectively. The intron discontinuity in *Isoetes* coincides with rearrangements in P9 which embraces a large continuous L9 loop of 800-nt in *Marchantia* (dotted line, blue). The *Marchantia* orthologue carries only traces of a formerly functional intron-encoded ORF. A C-to-U RNA editing event in the upstream exon is shown, three further cytidines in the intron core may likewise be subject to editings which could improve conservation of base-pairings.

both for the upstream part of *cox*1 extending downstream across several other genes and recombination points (*trn*S-R6-*trn*Y-*trn*I-R7-*sdh*3-*rpl*5-R10-*nad*1) as well as the downstream part of *cox*1 extending upstream across other genes (*trn*W-*nad*3-*rps*4), respectively (Figure 4B). Most notably, the different genomic routes downstream of *cox*1/5′ (Figure 2) identified through fosmid mapping were found to be faithfully reflected by two PCR products confirming the coexisting gene arrangements (Figure 4B).

RT–PCR products across the cox1i1305 discontinuity were easily retrieved from cDNA (Figure 4A). Cloning and sequencing verified correct splicing of cox1i1305 and all five additional upstream *cox*1 introns and showed differences to the genomic sequence exclusively at 106 positions of RNA editing, exactly as expected. Modelling the discontinuous group I intron sequences of cox1i1305 flanking the distantly located terminal *cox*1 exons reveals that the two sequence halves can combine for a classic group I secondary structure (Figure 5), to our knowledge the first example of a *trans*-splicing group I intron identified in nature. The typical ribozymic intron core structure of group I introns (40–42) is well conserved in comparison to its conventionally *cis*-arranged homologue in *Marchantia* (43).

To complement the PCR approaches outlined above in targeting potential alternatively arranged *cox*1 loci, we used *cox*1 cDNA as well as a mixture of the cox1i1305 intron halves as new probes. Rehybridization into our fosmid library, however, identified only those fosmids that had been identified and sequenced before.

To independently investigate the *cox*1 gene arrangement in *I. engelmannii* as deduced from the mtDNA map (Figure 2), we have used probes covering the terminal and sub-terminal exons 6 and 7 of the *cox*1 gene in a Southern blot hybridization experiment (Figure 6). Restriction sites for digestion of total genomic *Isoetes* DNA were selected to include the proximal recombination points identified near the upstream (R6) and the downstream part of the *cox*1 gene (R17). With hybridizations using probes for the upstream (Figure 6A) and downstream part of the *cox*1 gene (Figure 6B), two hybridizing restriction fragments were indeed identified in each case, reflecting the co-existing genomic rearrangements exactly as predicted from the genomic map (Figure 2). No further, additional hybridization signals were identified, which

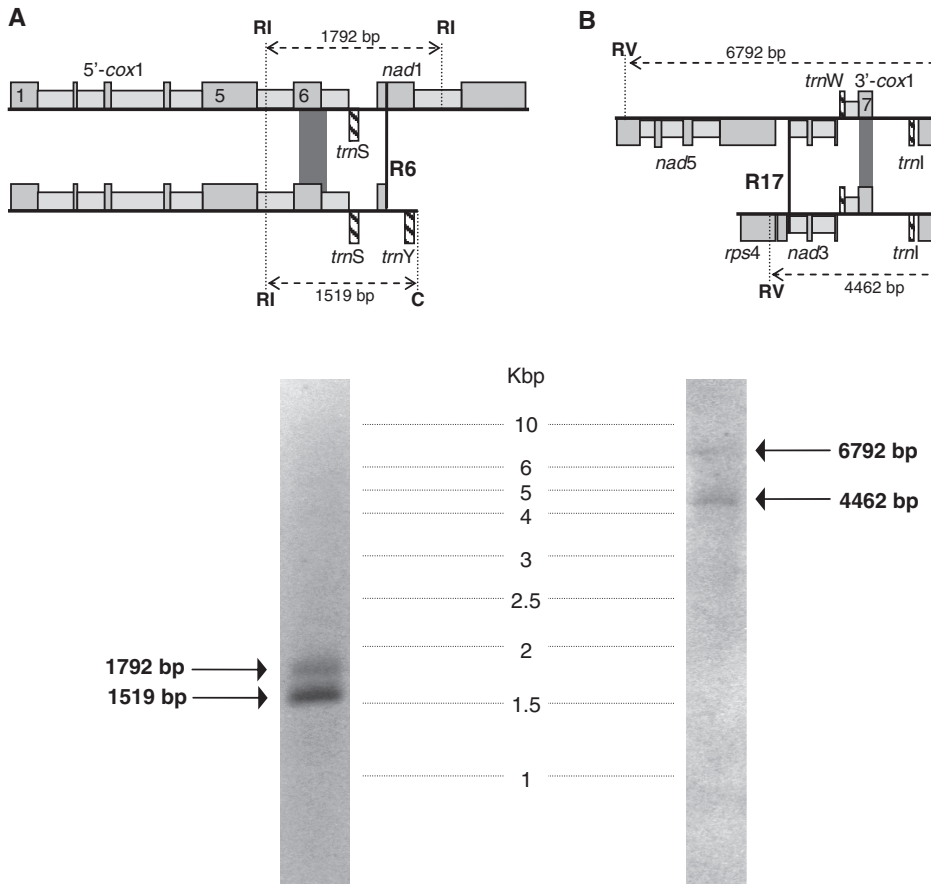

Figure 6. Southern-blot hybridization to verify *cox*1 gene arrangements for the upstream (A) and downstream (B) part of the gene using probes covering *cox*1 exons 6 and 7, respectively (dark grey rectangles). *Isoetes engelmannii* total genomic DNA was digested with *Eco*RI (RI) and *Cfr*9I (C) or with *Eco*RI (RI) and *Eco*RV (RV) to include the nearest identified recombination sites R6 or R17, respectively, in each case. Only two restriction fragments of expected sizes for coexisting genomic arrangements across R6 and R17 were identified by hybridization in each case: the 5′ *cox*1 part followed by *trn*S and either *nad*1 or by Ψ*nad*1-*trn*Y (**A**) and the 3′ *cox*1 part preceded by either *nad*5, Ψ*nad*3 and *trn*W or by *rps*4, *nad*3 and *trn*W (**B**).

could potentially represent a continuous *cox*1 gene copy (either intron-less or with a *cis*-splicing cox1i1305 counterpart) in full accord with the PCR experiments detailed above.

To determine *cox*1 transcript ends we used an approach of RNA circularization by self-ligation, followed by cDNA synthesis and RT–PCR with outward directed primers (oligonucleotides 1 and 2 in Figure 2). This revealed a 3′-UTR (untranslated region) extending 18 bp downstream of the stop codon and a 5′-UTR of 71 bp. The first seven *cox*1 codons are identical to those of the *atp*8 gene provided via recombination event R12 (Figure 2).

DISCUSSION

Lycophytes occupy a crucial position in the phylogeny of land plants (Figure 1), now unequivocally recognized as the sister group to euphyllophytes, which comprise the seed plants and the monilophytes with the latter encompassing the ferns, horsetails and whisk ferns (10,44,45). Comparatively poor in numbers of genera, families and with only three orders (Isoetales, Lycopodiales and Selaginellales) the recent lycophytes represent the most ancient lineage of vascular plants. As such, they could be expected to assume an intermediary position between the non-vascular bryophytes and the evolutionary advanced tracheophytes also with respect to the evolution of complexity in plant mtDNA. However, the *I. engelmannii* mitochondrial genome reported here as a first lycophyte mtDNA rather underlines the notion that plant mitochondria are 'more unique than ever' (46) by providing yet another example of unique pathways of organelle genome evolution.

Two main evolutionary trends have evidently shaped the *I. engelmannii* mtDNA. The gain and rise of recombinational activity seems to be the evolutionary force producing co-existing gene arrangements and the discontinuous group I intron now discovered in the *cox*1 gene. Likewise, highly active DNA recombination may be the ultimate prerequisite for the incorporation of DNA from the nuclear and chloroplast genomes, which has not been observed in bryophyte mtDNAs. After the recent report

of chloroplast DNA inserts in the mtDNA of the gymnosperm *Cycas taitungensis* (47), the first occurrences of such 'promiscuous' inserts of foreign DNA are now pushed back yet way further in plant evolution. The peculiar disposition of plant mtDNA to incorporate foreign genetic material originating from the other two genomes in the plant cell may have evolved with the increase of recombinational activity in the earliest tracheophytes (Figure 1).

The small introns and the small intergenic regions in *Isoetes* mtDNA on the other hand seem to reflect a counter-acting trend for organelle genome compaction. This tendency is also reflected in the *Isoetes* mtDNA gene complement itself, which generally mirrors the observations made for independent nuclear gene transfer in a survey of 280 flowering plant genera (38). Ribosomal protein genes that were found frequently and independently lost from the angiosperm mtDNAs are similarly missing from the *Isoetes* mtDNA, whereas those found to be transferred to the nucleus more rarely are (still) present (notably *rps*2, *rps*3 and *rps*4). Obvious exceptions on the other hand are the *ccm* genes not present in the *Isoetes* mtDNA. The inability to identify *ccm* genes independently via PCR may either indicate the commonly observed significant sequence alteration after nuclear gene transfer or, as a more remote possibility, an evolutionary switch to an alternative pathway of cytochrome *c* maturation (48,49). *Vice versa*, the *sdh*3 gene was here identified in the *Isoetes* mtDNA but is frequently transferred to the nucleus in angiosperms.

Notably, despite high recombinational activity, none of the 27 group II introns in the *I. engelmannii* mtDNA was found in a *trans*-splicing arrangement. On the contrary, four of the conserved *trans*-splicing group II introns of angiosperms find their small orthologues as *cis*-arranged counterparts in *Isoetes* (Table 1). Hence, the *trans*-splicing group I intron reported here to occur in the *cox*1 gene may represent a mere chance product with recombination acting before size reduction towards a minimum ribozyme core had reduced the chances of creating a discontinuous, yet functional, intron. *Trans*-splicing group II introns are known for more than 20 yrs since their discoveries both in chloroplasts (50,51) of algae and land plants and briefly thereafter in plant mitochondria (52–54) and, more recently, also in the mtDNA of an alga (55).

The first example of a *trans*-splicing group I Intron in nature shows that discontinuous molecules exist in yet another class of ribozyme-type RNAs after a discontinuous hammerhead RNA had been reported very recently (56). For mitochondrial genomes, yet another type of gene discontinuity with 'modules' distributed over separate DNA molecules for which the mechanisms of RNA maturation still have to be determined had recently been described for the protist *Diplonema* (57). Other examples for unusual modes of RNA maturation have also been reported outside of mitochondria such as tRNAs encoded in separate genes for 5′ and 3′ halves in *Nanoarchaeum* (58,59) or circularly permuted tRNAs expressed via circular RNA intermediates in the red alga *Cyanidioschyon* (60).

Whereas several of the introns in the *Isoetes* mtDNA have clear homologues at identical positions in bryophyte or seed plant mitochondrial genomes (Table 1), nine of the group II insertion sites are so far unique in the quillwort. It will be highly interesting whether homologues of these introns can be identified in the remaining major land plant clades for which complete mitochondrial genomes are still missing (Figure 1): ferns, horsetails, whisk ferns and hornworts given that their gains and losses could add independent further data relevant to the backbone of land plant phylogeny.

The extreme requirement of RNA editing in *I. engelmannii* mitochondrial RNAs not only to re-establish conserved reading frames in mRNAs but also to reconstitute secondary structures of tRNAs exceeds what has been observed before including the recent estimates of some 1000 editing sites in *Cycas taitungensis* mitochondrial mRNAs (32). The C-to-U editings hitherto observed in tRNAs of plant organelles were considered functionally essential, yet rare, events, e.g. (61–63). Likewise, such events of C-to-U editing have also been reported in animal mitochondria, e.g. (64). Multiple sites of RNA editing in single mitochondrial tRNA species had been observed as different types of editing: the replacement of nucleotides in tRNA acceptor stems, e.g. in the protist *Acanthamoeba castellanii* (65), the chytridomycete fungus *Spizellomyces punctatus* (66), de novo synthesis of 3′ ends in a centipede (67) or the pyrimidine insertional type of editing in slime molds (68).

Given the extraordinary degree of recombination, the presence of chloroplast and nuclear sequence inserts, a *trans*-splicing group I intron, the extraordinary amounts of RNA editing in mRNAs and, most notably, in hitherto unseen amounts also in tRNAs the *I. engelmannii* mtDNA once more demonstrates that 'anything goes' in mitochondrial genome evolution (69).

ACCESSION NUMBERS

FJ010859, FJ536259, FJ390841, FJ176330, FJ628360.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are highly grateful to Dr Jeffrey Palmer and co-workers (Bloomington, IN, USA.) for generously making fresh material of *Isoetes engelmannii* available to us, to Diana Kühn at the Max-Planck-Institut (MPI) für Züchtungsforschung, Cologne, for fosmid clone sorting and filter spotting. We also wish to thank Monika Polsakiewicz for her skillful technical assistance; and Julia Neuwirt, Yesim Kümetepe and Patrick Johner for their early work on lycophyte mtDNA cosmid cloning attempts. Finally, we are very grateful for helpful comments on this manuscript by two anonymous reviewers. Sequences obtained in this study were deposited in

GenBank (accession numbers FJ010859, FJ536259, FJ390841, FJ176330, FJ628360).

FUNDING

Deutsche Forschungsgemeinschaft [DFG Kn411/6-1].

Conflict of interest statement. None declared.

REFERENCES

1. Gray,M.W., Burger,G. and Lang,B.F. (1999) Mitochondrial evolution. *Science*, **283**, 1476–1481.
2. Lavrov,D.V. (2007) Key transitions in animal evolution: a mitochondrial DNA perspective. *Integr. Compar. Biol.*, **47**, 734–743.
3. Gray,M.W., Lang,B.F. and Burger,G. (2004) Mitochondria of protists. *Annu. Rev. Genet.*, **38**, 477–524.
4. Lang,B.F., Burger,G., O'Kelly,C.J., Cedergren,R., Golding,G.B., Lemieux,C., Sankoff,D., Turmel,M. and Gray,M.W. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature*, **387**, 493–497.
5. Vaidya,A.B., Akella,R. and Suplick,K. (1989) Sequences similar to genes for two mitochondrial proteins and portions of ribosomal RNA in tandemly arrayed 6-kilobase-pair DNA of a malarial parasite. *Mol. Biochem. Parasitol.*, **35**, 97–107.
6. Ward,B.L., Anderson,R.S. and Bendich,A.J. (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell*, **25**, 793–803.
7. Beagley,C.T., Okada,N.A. and Wolstenholme,D.R. (1996) Two mitochondrial group I introns in a metazoan, the sea anemone Metridium senile: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. *Proc. Natl Acad. Sci. USA*, **93**, 5619–5623.
8. Turmel,M., Otis,C. and Lemieux,C. (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. *Proc. Natl Acad. Sci. USA*, **99**, 11275–11280.
9. Turmel,M., Otis,C. and Lemieux,C. (2003) The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell*, **15**, 1888–1903.
10. Qiu,Y.L., Li,L., Wang,B., Chen,Z., Knoop,V., Groth-Malonek,M., Dombrovska,O., Lee,J., Kent,L., Rest,J. *et al.* (2006) The deepest divergences in land plants inferred from phylogenomic evidence. *Proc. Natl Acad. Sci. USA*, **103**, 15511–15516.
11. Stern,D.B. and Lonsdale,D.M. (1982) Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature*, **299**, 698–702.
12. Knoop,V., Unseld,M., Marienfeld,J., Brandt,P., Sünkel,S., Ullrich,H. and Brennicke,A. (1996) *copia*-, *gypsy*- and LINE-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. *Genetics*, **142**, 579–585.
13. Vaughn,J.C., Mason,M.T., Sper-Whitis,G.L., Kuhlman,P. and Palmer,J.D. (1995) Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric CoxI gene of Peperomia. *J. Mol. Evol.*, **41**, 563–572.
14. Cho,Y.R. and Palmer,J.D. (1999) Multiple acquisitions via horizontal transfer of a group I intron in the mitochondrial cox1 gene during evolution of the Araceae family. *Mol. Biol. Evol.*, **16**, 1155–1165.
15. Bergthorsson,U., Adams,K.L., Thomason,B. and Palmer,J.D. (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*, **424**, 197–201.
16. Richardson,A.O. and Palmer,J.D. (2006) Horizontal gene transfer in plants. *J. Exp. Bot.*, **58**, 1–9.
17. Kubo,T. and Mikami,T. (2007) Organization and variation of angiosperm mitochondrial genome. *Phys. Plant.*, **129**, 6–13.
18. Knoop,V. (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.*, **46**, 123–139.
19. Takenaka,M., Verbitskiy,D., van der Merwe,J.A., Zehrmann,A. and Brennicke,A. (2008) The process of RNA editing in plant mitochondria. *Mitochondrion*, **8**, 35–46.
20. Maier,R.M., Zeltz,P., Kössel,H., Bonnard,G., Gualberto,J.M. and Grienenberger,J.M. (1996) RNA editing in plant mitochondria and chloroplasts. *Plant Mol. Biol.*, **32**, 343–365.
21. Steinhauser,S., Beckert,S., Capesius,I., Malek,O. and Knoop,V. (1999) Plant mitochondrial RNA editing: extreme in hornworts and dividing the liverworts? *J. Mol. Evol.*, **48**, 303–312.
22. Oda,K., Yamato,K., Ohta,E., Nakamura,Y., Takemura,M., Nozato,N., Akashi,K., Kanegae,T., Ogura,Y., Kohchi,T. *et al.* (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J. Mol. Biol.*, **223**, 1–7.
23. Terasawa,K., Odahara,M., Kabeya,Y., Kikugawa,T., Sekine,Y., Fujiwara,M. and Sato,N. (2006) The mitochondrial genome of the moss *Physcomitrella patens* sheds new light on mitochondrial evolution in land plants. *Mol. Biol. Evol.*, **24**, 699–709.
24. Oldenburg,D.J. and Bendich,A.J. (2001) Mitochondrial DNA from the liverwort Marchantia polymorpha: circularly permuted linear molecules, head-to-tail concatemers, and a 5′ protein. *J. Mol. Biol.*, **310**, 549–562.
25. Unseld,M., Marienfeld,J.R., Brandt,P. and Brennicke,A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.*, **15**, 57–61.
26. Sugiyama,Y., Watase,Y., Nagase,M., Makita,N., Yagura,S., Hirai,A. and Sugiura,M. (2005) The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol. Genet. Genom.*, **272**, 603–615.
27. Ogihara,Y., Yamazaki,Y., Murai,K., Kanno,A., Terachi,T., Shiina,T., Miyashita,N., Nasuda,S., Nakamura,C., Mori,N. *et al.* (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res.*, **33**, 6235–6250.
28. Shedge,V., Arrieta-Montiel,M., Christensen,A.C. and Mackenzie,S.A. (2007) Plant mitochondrial recombination surveillance requires unusual RecA and MutS homologs. *Plant Cell*, **19**, 1251–1264.
29. Bendich,A.J. (1993) Reaching for the ring: the study of mitochondrial genome structure. *Curr. Genet.*, **24**, 279–290.
30. Malek,O. and Knoop,V. (1998) Trans-splicing group II introns in plant mitochondria: the complete set of *cis*-arranged homologs in ferns, fern allies, and a hornwort. *RNA*, **4**, 1599–1609.
31. Groth-Malonek,M., Pruchner,D., Grewe,F. and Knoop,V. (2005) Ancestors of *trans*-splicing mitochondrial introns support serial sister group relationships of hornworts and mosses with vascular plants. *Mol. Biol. Evol.*, **22**, 117–125.
32. Chaw,S.M., Chun-Chieh,S.A., Wang,D., Wu,Y.W., Liu,S.M. and Chou,T.Y. (2008) The mitochondrial genome of the gymnosperm *Cycas taitungensis* contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. *Mol. Biol. Evol.*, **25**, 603–615.
33. Lohse,M., Drechsel,O. and Bock,R. (2007) OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.*, **52**, 267–274.
34. Kumar,S., Nei,M., Dudley,J. and Tamura,K. (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.*, **9**, 299–306.
35. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
36. Kempken,F., Bolle,N., Forner,J. and Binder,S. (2008) Transcript end mapping and analysis of RNA editing in plant mitochondria. In Leister,D. and Herrmann,J.M. (eds), *Mitochondria: Practical Protocols,* Vol. 372, Humana Press, Totowa, NJ, pp. 177–192.
37. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
38. Adams,K.L., Qiu,Y.L., Stoutemyer,M. and Palmer,J.D. (2002) Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the

nucleus during angiosperm evolution. *Proc. Natl Acad. Sci. USA*, **99**, 9905–9912.
39. Dombrovska,E. and Qiu,Y.L. (2004) Distribution of introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implications. *Mol. Phylogenet. Evol.*, **32**, 246–263.
40. Lang,B.F., Laforest,M.J. and Burger,G. (2007) Mitochondrial introns: a critical view. *Trends Genet.*, **23**, 119–125.
41. Saldanha,R., Mohr,G., Belfort,M. and Lambowitz,A.M. (1993) Group I and group II introns. *FASEB J.*, **7**, 15–24.
42. Michel,F. and Westhof,E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.*, **216**, 585–610.
43. Ohta,E., Oda,K., Yamato,K., Nakamura,Y., Takemura,M., Nozato,N., Akashi,K., Ohyama,K. and Michel,F. (1993) Group I introns in the liverwort mitochondrial genome: the gene coding for subunit 1 of cytochrome oxidase shares five intron positions with its fungal counterparts. *Nucleic Acids Res.*, **21**, 1297–1305.
44. Pryer,K.M., Schneider,H., Smith,A.R., Cranfill,R., Wolf,P.G., Hunt,J.S. and Sipes,S.D. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature*, **409**, 618–622.
45. Qiu,Y.L. (2008) Phylogeny and evolution of charophytic algae and land plants. *J. Syst. Evol.*, **46**, 287–306.
46. Rasmusson,A.G., Handa,H. and Moller,I.M. (2008) Plant mitochondria, more unique than ever. *Mitochondrion*, **8**, 1–4.
47. Wang,D., Wu,Y.W., Shih,A.C., Wu,C.S., Wang,Y.N. and Chaw,S.M. (2007) Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. *Mol. Biol. Evol.*, **24**, 2040–2048.
48. Giegé,P., Grienenberger,J.M. and Bonnard,G. (2008) Cytochrome c biogenesis in mitochondria. *Mitochondrion*, **8**, 61–73.
49. Allen,J.W., Jackson,A.P., Rigden,D.J., Willis,A.C., Ferguson,S.J. and Ginger,M.L. (2008) Order within a mosaic distribution of mitochondrial c-type cytochrome biogenesis systems? *FEBS J.*, **275**, 2385–2402.
50. Koller,B., Fromm,H., Galun,E. and Edelman,M. (1987) Evidence for in vivo trans splicing of pre-mRNAs in tobacco chloroplasts. *Cell*, **48**, 111–119.
51. Choquet,Y., Goldschmidt-Clermont,M., Girard-Bascou,J., Kuck,U., Bennoun,P. and Rochaix,J.D. (1988) Mutant phenotypes support a trans-splicing mechanism for the expression of the tripartite psaA gene in the C. reinhardtii chloroplast. *Cell*, **52**, 903–913.
52. Knoop,V., Schuster,W., Wissinger,B. and Brennicke,A. (1991) *Trans* splicing integrates an exon of 22 nucleotides into the *nad5* mRNA in higher plant mitochondria. *EMBO J.*, **10**, 3483–3493.
53. Wissinger,B., Schuster,W. and Brennicke,A. (1991) *Trans* splicing in Oenothera mitochondria: *nad1* mRNAs are edited in exon and *trans*-splicing group II intron sequences. *Cell*, **65**, 473–482.
54. Chapdelaine,Y. and Bonen,L. (1991) The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. *Cell*, **65**, 465–472.
55. Turmel,M., Otis,C. and Lemieux,C. (2002) The complete mitochondrial DNA sequence of Mesostigma viride identifies this green alga as the earliest green plant divergence and predicts a highly compact mitochondrial genome in the ancestor of all green plants. *Mol. Biol. Evol.*, **19**, 24–38.
56. Martick,M., Horan,L.H., Noller,H.F. and Scott,W.G. (2008) A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA. *Nature*, **454**, 899–902.
57. Marande,W. and Burger,G. (2007) Mitochondrial DNA as a genomic jigsaw puzzle. *Science*, **318**, 415.
58. Randau,L., Münch,R., Hohn,M.J., Jahn,D. and Söll,D. (2005) *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5′- and 3′-halves. *Nature*, **433**, 537–541.
59. Randau,L. and Soll,D. (2008) Transfer RNA genes in pieces. *EMBO Rep.*, **9**, 623–628.
60. Soma,A., Onodera,A., Sugahara,J., Kanai,A., Yachie,N., Tomita,M., Kawamura,F. and Sekine,Y. (2007) Permuted tRNA genes expressed via a circular RNA intermediate in *Cyanidioschyzon merolae*. *Science*, **318**, 450–453.
61. Binder,S., Marchfelder,A. and Brennicke,A. (1994) RNA editing of $tRNA^{Phe}$ and $tRNA^{Cys}$ in mitochondria of *Oenothera berteriana* is initiated in precursor molecules. *Mol. Gen. Genet.*, **244**, 67–74.
62. Maréchal-Drouard,L., Ramamonjisoa,D., Cosset,A., Weil,J.H. and Dietrich,A. (1993) Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. *Nucleic Acids Res.*, **21**, 4909–4914.
63. Miyata,Y., Sugita,C., Maruyama,K. and Sugita,M. (2008) RNA editing in the anticodon of tRNA(Leu) (CAA) occurs before group I intron splicing in plastids of a moss *Takakia lepidozioides* S. Hatt. & Inoue. *Plant Biol (Stuttg)*, **10**, 250–255.
64. Janke,A. and Pääbo,S. (1993) Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Res.*, **21**, 1523–1525.
65. Lonergan,K.M. and Gray,M.W. (1993) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Science*, **259**, 812–816.
66. Laforest,M.J., Roewer,I. and Lang,B.F. (1997) Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. *Nucleic Acids Res.*, **25**, 626–632.
67. Lavrov,D.V., Brown,W.M. and Boore,J.L. (2000) A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl Acad. Sci USA*, **97**, 13738–13742.
68. Antes,T., Costandy,H., Mahendran,R., Spottswood,M. and Miller,D. (1998) Insertional editing of mitochondrial tRNAs of *Physarum polycephalum* and *Didymium nigripes*. *Mol. Cell Biol.*, **18**, 7521–7527.
69. Burger,G., Gray,M.W. and Lang,B.F. (2003) Mitochondrial genomes: anything goes. *Trends Genet.*, **19**, 709–716.